

RAW SEQUENCE LISTING

DATE: 09/29/2000

PATENT APPLICATION: US/09/479,467A

TIME: 18:28:27

Input Set : A:\Pto.amc

Output Set: N:\CRF3\09292000\I479467A.raw

```

3 <110> APPLICANT: Sternberg, Paul W.
4   Barr, Maureen M.
6 <120> TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE HOMOLOGS REQUIRED FOR MALE MATING
7   BEHAVIOR IN NEMATODES AND ASSAYS BASED THEREON
9 <130> FILE REFERENCE: 18021-2901B
11 <140> CURRENT APPLICATION NUMBER: 09/479,467A
12 <141> CURRENT FILING DATE: 2000-01-06
14 <150> PRIOR APPLICATION NUMBER: 60/115,127
15 <151> PRIOR FILING DATE: 1999-01-06
17 <160> NUMBER OF SEQ ID NOS: 16
19 <170> SOFTWARE: PatentIn Ver. 2.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 12912
23 <212> TYPE: DNA
24 <213> ORGANISM: Homo sapiens PKD-1 gene
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (1)..(12912)
30 <400> SEQUENCE: 1
31 atg ccg ccc gcc gcg ccc gcc cgc ctg gcg ctg gcc ctg gcc ctg gcc 48
32 Met Pro Pro Ala Ala Pro Ala Arg Leu Ala Leu Ala Leu Gly Leu Gly
33 1 5 10 15
35 ctg tgg ctc ggg gcg ctg gcg ggg ggg ccc ggc cgc ggc tgc ggg ccc 96
36 Leu Trp Leu Gly Ala Leu Ala Gly Gly Pro Gly Arg Gly Cys Gly Pro
37 20 25 30
39 tgc gag ccc ccc tgc ctc tgc ggg cca gcg ccc gcc gcc gcc tgc cgc 144
40 Cys Glu Pro Pro Cys Leu Cys Gly Pro Ala Pro Gly Ala Ala Cys Arg
41 35 40 45
43 gtc aac tgc tcg gcc cgc ggg ctg cgg acg ctc ggt ccc gcg ctg cgc 192
44 Val Asn Cys Ser Gly Arg Gly Leu Arg Thr Leu Gly Pro Ala Leu Arg
45 50 55 60
47 atc ccc gcg gac gcc aca gag cta gac gtc tcc cac aac ctg ctc cgg 240
48 Ile Pro Ala Asp Ala Thr Glu Leu Asp Val Ser His Asn Leu Leu Arg
49 65 70 75 80
51 gcg ctg gac gtt ggg ctc ctg gcg aac ctc tcg gcg ctg gca gag ctg 288
52 Ala Leu Asp Val Gly Leu Leu Ala Asn Leu Ser Ala Leu Ala Glu Leu
53 85 90 95
55 gat ata agc aac aac aag att tct acg tta gaa gaa gga ata ttt gct 336
56 Asp Ile Ser Asn Asn Lys Ile Ser Thr Leu Glu Glu Gly Ile Phe Ala
57 100 105 110
59 aat tta ttt aat tta agt gaa ata aac ctg agt ggg aac ccg ttt gag 384
60 Asn Leu Phe Asn Leu Ser Glu Ile Asn Leu Ser Gly Asn Pro Phe Glu
61 115 120 125
63 tgt gac tgt gcc ctg gcg tgg ctg ccg caa tgg gcg gag gag cag cag 432
64 Cys Asp Cys Gly Leu Ala Trp Leu Pro Gln Trp Ala Glu Glu Gln Gln
65 130 135 140
67 gtg cgg gtg gtg cag ccc gag gca gcc acg tgt gct ggg cct gcc tcc 480

```

RAW SEQUENCE LISTING

DATE: 09/29/2000

PATENT APPLICATION: US/09/479,467A

TIME: 18:28:27

Input Set : A:\Pto.amc

Output Set: N:\CRF3\09292000\I479467A.raw

```

68 Val Arg Val Val Gln Pro Glu Ala Ala Thr Cys Ala Gly Pro Gly Ser
69 145 150 155 160
71 ctg gct ggc cag cct ctg ctt ggc atc ccc ttg ctg gac agt ggc tgt 528
72 Leu Ala Gly Gln Pro Leu Leu Gly Ile Pro Leu Leu Asp Ser Gly Cys
73 165 170 175
75 ggt gag gag tat gtc gcc tgc ctc cct gac aac agc tca ggc acc gtg 576
76 Gly Glu Glu Tyr Val Ala Cys Leu Pro Asp Asn Ser Ser Gly Thr Val
77 180 185 190
79 gca gca gtg tcc ttt tca gct gcc cac gaa ggc ctg ctt cag cca gag 624
80 Ala Ala Val Ser Phe Ser Ala Ala His Glu Gly Leu Leu Gln Pro Glu
81 195 200 205
83 gcc tgc agc gcc ttc tgc ttc acc ggc cag ggc ctc gca gcc ctc 672
84 Ala Cys Ser Ala Phe Cys Phe Ser Thr Gly Gln Gly Leu Ala Ala Leu
85 210 215 220
87 tcg gag cag ggc tgg tgc ctg tgt ggg gcg gcc cag ccc tcc agt gcc 720
88 Ser Glu Gln Gly Trp Cys Leu Cys Gly Ala Ala Gln Pro Ser Ser Ala
89 225 230 235 240
91 tcc ttt gcc tgc ctg tcc ctc tgc tcc ggg ccc ccg gca cct cct gcc 768
92 Ser Phe Ala Cys Leu Ser Leu Cys Ser Gly Pro Pro Ala Pro Pro Ala
93 245 250 255
95 ccc acc tgt agg ggc ccc acc ctc ctc cag cac gtc ttc cct gcc tcc 816
96 Pro Thr Cys Arg Gly Pro Thr Leu Leu Gln His Val Phe Pro Ala Ser
97 260 265 270
99 cca ggg gcc acc ctg gtg ggg ccc cac gga cct ctg gcc tct ggc cag 864
100 Pro Gly Ala Thr Leu Val Gly Pro His Gly Pro Leu Ala Ser Gly Gln
101 275 280 285
103 cta gca gcc ttc cac atc gct gcc ccg ctc cct gtc act gac aca cgc 912
104 Leu Ala Ala Phe His Ile Ala Ala Pro Leu Pro Val Thr Asp Thr Arg
105 290 295 300
107 tgg gac ttc gga gac ggc tcc gcc gag gtg gat gcc gct ggg ccg gct 960
108 Trp Asp Phe Gly Asp Gly Ser Ala Glu Val Asp Ala Ala Gly Pro Ala
109 305 310 315 320
111 gcc tcg cat cgc tat gtg ctg cct ggg cgc tat cac gtg acg gcc gtg 1008
112 Ala Ser His Arg Tyr Val Leu Pro Gly Arg Tyr His Val Thr Ala Val
113 325 330 335
115 ctg gcc ctg ggg gcc ggc tca gcc ctg ctg ggg aca gac gtg cag gtg 1056
116 Leu Ala Leu Gly Ala Gly Ser Ala Leu Leu Gly Thr Asp Val Gln Val
117 340 345 350
119 gaa gcg gca cct gcc gcc ctg gag ctc gtg tgc ccg tcc tcg gtg cag 1104
120 Glu Ala Ala Pro Ala Ala Leu Glu Leu Val Cys Pro Ser Ser Val Gln
121 355 360 365
123 agt gac gag agc ctc gac ctc agc atc cag aac cgc ggt ggt tca ggc 1152
124 Ser Asp Glu Ser Leu Asp Leu Ser Ile Gln Asn Arg Gly Gly Ser Gly
125 370 375 380
127 ctg gag gcc gcc tac agc atc gtg gcc ctg ggc gag gag ccg gcc cga 1200
128 Leu Glu Ala Ala Tyr Ser Ile Val Ala Leu Gly Glu Glu Pro Ala Arg
129 385 390 395 400
131 gcg gtg cac ccg ctc tgc ccc tcg gac acg gag atc ttc cct ggc aac 1248
132 Ala Val His Pro Leu Cys Pro Ser Asp Thr Glu Ile Phe Pro Gly Asn

```

RAW SEQUENCE LISTING

DATE: 09/29/2000

PATENT APPLICATION: US/09/479,467A

TIME: 18:28:27

Input Set : A:\Pto.amc

Output Set: N:\CRF3\09292000\I479467A.raw

133				405				410				415					
135	ggg	cac	tgc	tac	cgc	ctg	gtg	gtg	gag	aag	gcg	gcc	tgg	ctg	cag	gcg	1296
136	Gly	His	Cys	Tyr	Arg	Leu	Val	Val	Glu	Lys	Ala	Ala	Trp	Leu	Gln	Ala	
137				420				425					430				
139	cag	gag	cag	tgt	cag	gcc	tgg	gcc	ggg	gcc	gcc	ctg	gca	atg	gtg	gac	1344
140	Gln	Glu	Gln	Cys	Gln	Ala	Trp	Ala	Gly	Ala	Ala	Leu	Ala	Met	Val	Asp	
141				435				440					445				
143	agt	ccc	gcc	gtg	cag	cgc	ttc	ctg	gtc	tcc	cgg	gtc	acc	agg	agc	cta	1392
144	Ser	Pro	Ala	Val	Gln	Arg	Phe	Leu	Val	Ser	Arg	Val	Thr	Arg	Ser	Leu	
145				450				455					460				
147	gac	gtg	tgg	atc	ggc	ttc	tcg	act	gtg	cag	ggg	gtg	gag	gtg	ggc	cca	1440
148	Asp	Val	Trp	Ile	Gly	Phe	Ser	Thr	Val	Gln	Gly	Val	Glu	Val	Gly	Pro	
149	465				470				475				480				
151	gcg	ccg	cag	ggc	gag	gcc	ttc	agc	ctg	gag	agc	tgc	cag	aac	tgg	ctg	1488
152	Ala	Pro	Gln	Gly	Glu	Ala	Phe	Ser	Leu	Glu	Ser	Cys	Gln	Asn	Trp	Leu	
153				485					490				495				
155	ccc	ggg	gag	cca	cac	cca	gcc	aca	gcc	gag	cac	tgc	gtc	cgg	ctc	ggg	1536
156	Pro	Gly	Glu	Pro	His	Pro	Ala	Thr	Ala	Glu	His	Cys	Val	Arg	Leu	Gly	
157				500					505				510				
159	ccc	acc	ggg	tgg	tgt	aac	acc	gac	ctg	tgc	tca	gcg	ccg	cac	agc	tac	1584
160	Pro	Thr	Gly	Trp	Cys	Asn	Thr	Asp	Leu	Cys	Ser	Ala	Pro	His	Ser	Tyr	
161				515					520				525				
163	gtc	tgc	gag	ctg	cag	ccc	gga	ggc	cca	gtg	cag	gat	gcc	gag	aac	ctc	1632
164	Val	Cys	Glu	Leu	Gln	Pro	Gly	Gly	Pro	Val	Gln	Asp	Ala	Glu	Asn	Leu	
165				530					535				540				
167	ctc	gtg	gga	gcg	ccc	agt	ggg	gac	ctg	cag	gga	ccc	ctg	acg	cct	ctg	1680
168	Leu	Val	Gly	Ala	Pro	Ser	Gly	Asp	Leu	Gln	Gly	Pro	Leu	Thr	Pro	Leu	
169	545				550				555				560				
171	gca	cag	cag	gac	ggc	ctc	tca	gcc	ccg	cac	gag	ccc	gtg	gag	gtc	atg	1728
172	Ala	Gln	Gln	Asp	Gly	Leu	Ser	Ala	Pro	His	Glu	Pro	Val	Glu	Val	Met	
173				565					570				575				
175	gta	ttc	ccg	ggc	ctg	cgt	ctg	agc	cgt	gaa	gcc	ttc	ctc	acc	acg	gcc	1776
176	Val	Phe	Pro	Gly	Leu	Arg	Leu	Ser	Arg	Glu	Ala	Phe	Leu	Thr	Thr	Ala	
177				580					585				590				
179	gaa	ttt	ggg	acc	cag	gag	ctc	cgg	cgg	ccc	gcc	cag	ctg	cgg	ctg	cag	1824
180	Glu	Phe	Gly	Thr	Gln	Glu	Leu	Arg	Arg	Pro	Ala	Gln	Leu	Arg	Leu	Gln	
181				595					600				605				
183	gtg	tac	cgg	ctc	ctc	agc	aca	gca	ggg	acc	ccg	gag	aac	ggc	agc	gag	1872
184	Val	Tyr	Arg	Leu	Leu	Ser	Thr	Ala	Gly	Thr	Pro	Glu	Asn	Gly	Ser	Glu	
185				610					615				620				
187	cct	gag	agc	agg	tcc	ccg	gac	aac	agg	acc	cag	ctg	gcc	ccc	gcg	tgc	1920
188	Pro	Glu	Ser	Arg	Ser	Pro	Asp	Asn	Arg	Thr	Gln	Leu	Ala	Pro	Ala	Cys	
189	625				630						635				640		
191	atg	cca	ggg	gga	cgc	tgg	tgc	cct	gga	gcc	aac	atc	tgc	ttg	ccg	ctg	1968
192	Met	Pro	Gly	Gly	Arg	Trp	Cys	Pro	Gly	Ala	Asn	Ile	Cys	Leu	Pro	Leu	
193				645					650				655				
195	gac	gcc	tcc	tgc	cac	ccc	cag	gcc	tgc	gcc	aat	ggc	tgc	acg	tca	ggg	2016
196	Asp	Ala	Ser	Cys	His	Pro	Gln	Ala	Cys	Ala	Asn	Gly	Cys	Thr	Ser	Gly	
197				660					665				670				

RAW SEQUENCE LISTING

DATE: 09/29/2000

PATENT APPLICATION: US/09/479,467A

TIME: 18:28:27

Input Set : A:\Pto.amc

Output Set: N:\CRF3\09292000\I479467A.raw

199	cca ggg cta ccc ggg gcc ccc tat gcg cta tgg aga gag ttc ctc ttc	2064
200	Pro Gly Leu Pro Gly Ala Pro Tyr Ala Leu Trp Arg Glu Phe Leu Phe	
201	675 680 685	
203	tcc gtt ccc gcg ggg ccc ccc gcg cag tac tcg gtc acc ctc cac ggc	2112
204	Ser Val Pro Ala Gly Pro Pro Ala Gln Tyr Ser Val Thr Leu His Gly	
205	690 695 700	
207	cag gat gtc ctc atg ctc cct ggt gac ctc gtt ggc ttg cag cac gac	2160
208	Gln Asp Val Leu Met Leu Pro Gly Asp Leu Val Gly Leu Gln His Asp	
209	705 710 715 720	
211	gct ggc cct ggc gcc ctc ctg cac tgc tcg ccg gct ccc ggc cac cct	2208
212	Ala Gly Pro Gly Ala Leu Leu His Cys Ser Pro Ala Pro Gly His Pro	
213	725 730 735	
215	ggt ccc cgg gcc ccg tac ctc tcc gcc aac gcc tcg tca tgg ctg ccc	2256
216	Gly Pro Arg Ala Pro Tyr Leu Ser Ala Asn Ala Ser Ser Trp Leu Pro	
217	740 745 750	
219	cac ttg cca gcc cag ctg gag ggc act tgg ggc tgc cct gcc tgt gcc	2304
220	His Leu Pro Ala Gln Leu Glu Gly Thr Trp Gly Cys Pro Ala Cys Ala	
221	755 760 765	
223	ctg cgg ctg ctt gca caa cgg gaa cag ctc acc gtg ctg ctg ggc ttg	2352
224	Leu Arg Leu Leu Ala Gln Arg Glu Gln Leu Thr Val Leu Leu Gly Leu	
225	770 775 780	
227	agg ccc aac cct gga ctg cgg ctg cct ggg cgc tat gag gtc cgg gca	2400
228	Arg Pro Asn Pro Gly Leu Arg Leu Pro Gly Arg Tyr Glu Val Arg Ala	
229	785 790 795 800	
231	gag gtg ggc aat ggc gtg tcc agg cac aac ctc tcc tgc agc ttt gac	2448
232	Glu Val Gly Asn Gly Val Ser Arg His Asn Leu Ser Cys Ser Phe Asp	
233	805 810 815	
235	gtg gtc tcc cca gtg gct ggg ctg cgg gtc atc tac cct gcc ccc cgc	2496
236	Val Val Ser Pro Val Ala Gly Leu Arg Val Ile Tyr Pro Ala Pro Arg	
237	820 825 830	
239	gac ggc cgc ctc tac gtg ccc acc aac ggc tca gcc ttg gtg ctc cag	2544
240	Asp Gly Arg Leu Tyr Val Pro Thr Asn Gly Ser Ala Leu Val Leu Gln	
241	835 840 845	
243	gtg gac tct ggt gcc aac gcc acg gcc acg gct cgc tgg cct ggg ggc	2592
244	Val Asp Ser Gly Ala Asn Ala Thr Ala Thr Ala Arg Trp Pro Gly Gly	
245	850 855 860	
247	agt ctc agc gcc cgc ttt gag aat gtc tgc cct gcc ctg gtg gcc acc	2640
248	Ser Leu Ser Ala Arg Phe Glu Asn Val Cys Pro Ala Leu Val Ala Thr	
249	865 870 875 880	
251	ttc gtg ccc gcc tgc ccc tgg gag acc aac gat acc ctg ttc tca gtg	2688
252	Phe Val Pro Ala Cys Pro Trp Glu Thr Asn Asp Thr Leu Phe Ser Val	
253	885 890 895	
255	gta gca ctg ccg tgg ctc agt gag ggg gag cac gtg gtg gac gtg gtg	2736
256	Val Ala Leu Pro Trp Leu Ser Glu Gly Glu His Val Val Asp Val Val	
257	900 905 910	
259	gtg gaa aac agc gcc agc cgg gcc aac ctc agc ctg cgg gtg acg gcg	2784
260	Val Glu Asn Ser Ala Ser Arg Ala Asn Leu Ser Leu Arg Val Thr Ala	
261	915 920 925	
263	gag gag ccc atc tgt ggc ctc cgc gcc acg ccc agc ccc gag gcc cgt	2832

RAW SEQUENCE LISTING

DATE: 09/29/2000

PATENT APPLICATION: US/09/479,467A

TIME: 18:28:27

Input Set : A:\Pto.amc

Output Set: N:\CRF3\09292000\I479467A.raw

264	Glu	Glu	Pro	Ile	Cys	Gly	Leu	Arg	Ala	Thr	Pro	Ser	Pro	Glu	Ala	Arg	
265		930					935				940						
267	gta	ctg	cag	gga	gtc	cta	gtg	agg	tac	agc	ccc	gtg	gtg	gag	gcc	ggc	2880
268	Val	Leu	Gln	Gly	Val	Leu	Val	Arg	Tyr	Ser	Pro	Val	Val	Glu	Ala	Gly	
269	945					950					955					960	
271	tcg	gac	atg	gtc	ttc	cgg	tgg	acc	atc	aac	gac	aag	cag	tcc	ctg	acc	2928
272	Ser	Asp	Met	Val	Phe	Arg	Trp	Thr	Ile	Asn	Asp	Lys	Gln	Ser	Leu	Thr	
273					965					970					975		
275	ttc	cag	aac	gtg	gtc	ttc	aat	gtc	att	tat	cag	agc	gcg	gcg	gtc	ttc	2976
276	Phe	Gln	Asn	Val	Val	Phe	Asn	Val	Ile	Tyr	Gln	Ser	Ala	Ala	Val	Phe	
277					980					985					990		
279	aag	ctc	tca	ctg	acg	gcc	tcc	aac	cac	gtg	agc	aac	gtc	acc	gtg	aac	3024
280	Lys	Leu	Ser	Leu	Thr	Ala	Ser	Asn	His	Val	Ser	Asn	Val	Thr	Val	Asn	
281			995				1000				1005						
283	tac	aac	gta	acc	gtg	gag	cgg	atg	aac	agg	atg	cag	ggt	ctg	cag	gtc	3072
284	Tyr	Asn	Val	Thr	Val	Glu	Arg	Met	Asn	Arg	Met	Gln	Gly	Leu	Gln	Val	
285		1010					1015				1020						
287	tcc	aca	gtg	ccg	gcc	gtg	ctg	tcc	ccc	aat	gcc	acg	cta	gca	ctg	acg	3120
288	Ser	Thr	Val	Pro	Ala	Val	Leu	Ser	Pro	Asn	Ala	Thr	Leu	Ala	Leu	Thr	
289		1025				1030				1035					1040		
291	gcg	ggc	gtg	ctg	gtg	gac	tcg	gcc	gtg	gag	gtg	gcc	ttc	ctg	tgg	acc	3168
292	Ala	Gly	Val	Leu	Val	Asp	Ser	Ala	Val	Glu	Val	Ala	Phe	Leu	Trp	Thr	
293					1045					1050					1055		
295	ttt	ggg	gat	ggg	gag	cag	gcc	ctc	cac	cag	ttc	cag	cct	ccg	tac	aac	3216
296	Phe	Gly	Asp	Gly	Glu	Gln	Ala	Leu	His	Gln	Phe	Gln	Pro	Pro	Tyr	Asn	
297			1060				1065				1070						
299	gag	tcc	ttc	cca	gtt	cca	gac	ccc	tcg	gtg	gcc	cag	gtg	ctg	gtg	gag	3264
300	Glu	Ser	Phe	Pro	Val	Pro	Asp	Pro	Ser	Val	Ala	Gln	Val	Leu	Val	Glu	
301			1075				1080				1085						
303	cac	aat	gtc	acg	cac	acc	tac	gct	gcc	cca	ggt	gag	tac	ctc	ctg	acc	3312
304	His	Asn	Val	Thr	His	Thr	Tyr	Ala	Ala	Pro	Gly	Glu	Tyr	Leu	Leu	Thr	
305		1090					1095				1100						
307	gtg	ctg	gca	tct	aat	gcc	ttc	gag	aac	ctg	acg	cag	cag	gtg	cct	gtg	3360
308	Val	Leu	Ala	Ser	Asn	Ala	Phe	Glu	Asn	Leu	Thr	Gln	Gln	Val	Pro	Val	
309	1105					1110					1115				1120		
311	agc	gtg	cgc	gcc	tcc	ctg	ccc	tcc	gtg	gct	gtg	ggt	gtg	agt	gac	ggc	3408
312	Ser	Val	Arg	Ala	Ser	Leu	Pro	Ser	Val	Ala	Val	Gly	Val	Ser	Asp	Gly	
313				1125					1130					1135			
315	gtc	ctg	gtg	gcc	ggc	cgg	ccc	gtc	acc	ttc	tac	ccg	cac	ccg	ctg	ccc	3456
316	Val	Leu	Val	Ala	Gly	Arg	Pro	Val	Thr	Phe	Tyr	Pro	His	Pro	Leu	Pro	
317			1140					1145					1150				
319	tcg	cct	ggg	ggt	gtt	ctt	tac	acg	tgg	gac	ttc	ggg	gac	ggc	tcc	cct	3504
320	Ser	Pro	Gly	Gly	Val	Leu	Tyr	Thr	Trp	Asp	Phe	Gly	Asp	Gly	Ser	Pro	
321		1155					1160				1165						
323	gtc	ctg	acc	cag	agc	cag	ccg	gct	gcc	aac	cac	acc	tat	gcc	tcg	agg	3552
324	Val	Leu	Thr	Gln	Ser	Gln	Pro	Ala	Ala	Asn	His	Thr	Tyr	Ala	Ser	Arg	
325		1170					1175				1180						
327	ggc	acc	tac	cac	gtg	cgc	ctg	gag	gtc	aac	aac	acg	gtg	agc	ggt	gcg	3600
328	Gly	Thr	Tyr	His	Val	Arg	Leu	Glu	Val	Asn	Asn	Thr	Val	Ser	Gly	Ala	

VERIFICATION SUMMARY

DATE: 09/29/2000

PATENT APPLICATION: US/09/479,467A

TIME: 18:28:28

Input Set : A:\Pto.amc

Output Set: N:\CRF3\09292000\I479467A.raw